AMENDMENTS TO THE CLAIMS

1. - 11. (Canceled)

12. (Currently Amended) A method performed by a suitably programmed computer for inferring genomic sequences unique to at least one set of organisms other than a set of organisms under investigation, the method comprising:

obtaining genomic data characteristic of a set of organisms under investigation;

formatting by a computer program embodied on a computer-readable medium the genomic data into at least one query-length sequence, each query-length sequence being of a format compatible with a similarity search engine program;

searching, by the similarity search engine program, a selected genomic computer database using the query-length sequence and the similarity search engine program for those sequences having homology above a threshold with at least one set of organisms other than the set under investigation,

the selected genomic computer database containing genomic data from a plurality of organisms;

parsing results of the search for a sequence unique to the selected genomic database;

identifying from the selected genomic database the results of the search that are not a unique sequence and re-evaluating a subset of those results for unique sequences

removing the sequences from the selected genomic database that are not unique until only evaluated sequences that are unique remain; and

outputting by the computer to a user an identity of those sequences unique to the selected genomic database.

13. (Currently Amended) A computer-implemented system for inferring genomic sequences unique to a at least one set of organisms other than a set of organisms under investigation, the computer-implemented system comprising:

a computer program embodied on a computer ~~computer-readable physical medium~~, the computer program comprising a genomic data interface module, a formatting

module, a search interface module, a search results parsing module, and a removal module;

the genomic data interface module on the computer couples to a source of genomic data to receive genomic data characteristic of a set of organisms under investigation;

the formatting module on the computer formats received genomic data into at least one query-length sequence, each query-length sequence being of a format compatible with a similarity search engine;

the search interface module on the computer interfaces with the similarity search engine to submit the query-length sequence to a selected genomic database containing genomic data from a substantial plurality of organisms;

the search results parsing module on the computer parses results of the search for those sequences having homology above a threshold with at least one set of organisms other than the set under investigation and otherwise unique within the selected genomic database, and outputs to a user an identity of those sequences having homology above a threshold with at least one set of organisms other than the set under investigation and otherwise unique within the selected genomic database; and

the removal module on the computer removes the sequences from the selected genomic database that are not unique until only evaluated sequences that are unique remain;

wherein a subset of the results of the search that are not a unique sequence are re-evaluated by the search interface module.

14. (Currently Amended) A computer-implemented ~~program, comprising a computer readable medium having a computer readable program code embodied therein, the computer readable program code adapted to be executed to implement a~~ method for inferring genomic sequences unique to a first set of organisms, the method comprising:

providing a system, wherein the system comprises distinct software modules, and wherein the distinct software modules comprise a formatting module, a similarity search engine module, and a parsing module;

obtaining genomic data characteristic of a second set of organisms;

formatting by a computer the second set genomic data into at least one query-length sequence, each query-length sequence being of a format compatible with the similarity search engine module, wherein the formatting is performed by the formatting module;

searching by the similarly search engine module a selected genomic database using the formatted query-length sequence, the selected genomic database containing genomic data from a plurality of organisms including the first set of organisms;

parsing results of the search for those sequences from the similarity search engine module, other than sequences of the second set, having homology above a threshold with the second set and otherwise unique within the selected genomic database, wherein the parsing is performed by the parsing module;

identifying from the selected genomic database the results of the search that are not a unique sequence and re-evaluating a subset of those results for unique sequences;

removing the sequences from the selected genomic database that are not unique until only evaluated sequences that are unique remain; and

outputting by the computer to a user an identity of those sequences having homology above a threshold with the second set and otherwise unique within the selected genomic database as genomic sequences unique to the first set.

15. (Canceled)